



Détections des émotions sur un smartphone à partir des interactions des utilisateurs. Proposition de protocole expérimental.

Nicolas Simonazzi, Jean-Marc Salotti, Caroline Dubois, Dominique Séminel

► To cite this version:

Nicolas Simonazzi, Jean-Marc Salotti, Caroline Dubois, Dominique Séminel. Détections des émotions sur un smartphone à partir des interactions des utilisateurs. Proposition de protocole expérimental.. 54ème Congrès International de la Société d'Ergonomie de Langue Française, Sep 2019, Tours, France. hal-02393182

HAL Id: hal-02393182

<https://hal.science/hal-02393182>

Submitted on 4 Dec 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Texte original*.

Détections des émotions sur un smartphone à partir des interactions des utilisateurs. Proposition de protocole expérimental.

N. Simonazzi

J.M. Salotti

C. Dubois

D. Séminel

ORANGE Labs – 4, rue du Clos Courtel 35510 Cesson-Sévigné
Equipe Cognitive Equipe Auctus, INRIA, Bordeaux INP, Laboratoire IMS, Université
de Bordeaux, CNRS, 109 Avenue Roul, 33400.

nicolas.simonazzi@ensc.fr
caroline.dubois@orange.com

jean-marc.salotti@ensc.fr
dominique.seminel@orange.com

Dans cette étude nous décrivons l'élaboration d'un dispositif de mesure des émotions, étayé par la revue des principales théories et recherches expérimentales déjà réalisées dans ce domaine. Pour cela, nous proposons d'exploiter les données comportementales recueillies à partir des capteurs du smartphone. Dans notre protocole expérimental, les participants seront ainsi amenés à interagir avec un agent conversationnel dans une application de chat. Les tâches et le contenu proposés ont pour objectif d'induire des émotions chez les participants pour labéliser les données rassemblées et construire un modèle de prédiction. Les perspectives que nous envisageons pour ces travaux sont relatives au développement d'outils susceptibles d'être utilisés dans le cadre de la pratique ergonomique.

Mots-clés : Emotions, Conception de systèmes et ingénierie d'interface, Dispositifs d'entrée tactiles, Méthodes d'analyse et de traitement des données.

Emotion detection through smartphone from users' interaction : proposal for experimental protocol.

With this research, we've outlined the development of a device intended to measure and quantify emotions. To back up our work, we have relied on the main theories and experimental research that have already been conducted in this particular field. In order to conduct this research, we've chosen to use behavioral data collected from a smartphone's user. During our experimental process, participants will be asked to interact with a chatbot application. Both tasks and contents are intended to induce emotions for the participants, so as to label collected data and build a prediction model. The prospects we are considering in this research are closely related to the development of tools which may be used in ergonomic practices.

Keywords: Emotion, System design and interface engineering, Tactile input devices, Data analysis and processing methods

*Ce texte original a été produit dans le cadre du congrès de la Société d'Ergonomie de Langue Française qui s'est tenu à Tours, les 25, 26 et 27 septembre 2019. Il est permis d'en faire une copie papier ou digitale pour un usage pédagogique ou universitaire, en citant la source exacte du document, qui est la suivante :

Détections des émotions sur un smartphone à partir des interactions des utilisateurs. Proposition d'un protocole expérimental.. Actes du 54^{ème} Congrès de la SELF, Université de l'Ergonomie : Comment contribuer à un autre monde ? Tours, 25, 26 et 27 septembre 2019

Aucun usage commercial ne peut en être fait sans l'accord des éditeurs ou archiveurs électroniques. Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page.

INTRODUCTION

Le cadre des services clients des opérateurs télécom connaît un contexte en pleine mutation, principalement au regard de l'introduction d'outils numériques de plus en plus nombreux. Afin de limiter les coûts de la relation client, les entreprises tendent à automatiser cette relation de service en déléguant les tâches les plus répétitives et ayant le moins de valeur ajoutée à des automates (Velkovska, 2014). Ainsi, l'explosion de l'utilisation des chatbots en est un exemple.

Les émotions font parties intégrantes de nos communications interhumaines, directes ou médiées par des machines. L'auteur Brave s'appuie sur la pyramide de Maslow (Brave, 2003), (Maslow, 1968) pour expliquer la place des émotions dans la communication. Ainsi, les clients/utilisateurs ont des attentes (sociales, cognitives..) auxquelles ils tentent de répondre lors des interactions de communication. Si ces attentes sont satisfaites, il est probable que l'utilisateur ressent une émotion positive et inversement si ce n'est pas le cas. La conception d'outils permettant de mesurer les affects des clients serait donc un axe de travaux à forte valeur pour les entreprises, l'émotion étant un facteur essentiel dans les processus de fidélisation. Pourtant, très peu de travaux portent actuellement sur la détection des émotions dans le cadre d'un service client.

Dans le cas d'une conversation par chat, nous proposons de détecter les émotions à partir de l'usage de smartphones. De fait, les smartphones sont très intensivement utilisés au quotidien. Ils constituent dès lors une plateforme très adaptée pour recueillir des données comportementales grâce aux différents capteurs dont ils sont dotés. Une question méthodologique importante émerge de ce choix : les données comportementales issues des smartphones peuvent-elles aider à détecter les émotions des utilisateurs ?

Nous proposons dans une première partie de faire l'état de l'art des modèles

émotionnels et de présenter plus en détail celui qui a été retenu pour notre étude. Dans une deuxième partie, nous décrivons le protocole expérimental que nous souhaitons mettre en œuvre pour tenter de détecter les émotions grâce aux capteurs des smartphones. Enfin, nous abordons les principaux axes de recherche qui feront suite à ces travaux et les liens avec l'ergonomie cognitive.

ETAT DE L'ART

Rôle des émotions

La compréhension des émotions dans un environnement de plus en plus numérique a été particulièrement défendue par Rosalind Picard qui a présenté et défini la discipline de l'informatique affective (Picard, 1995). En effet, Picard postule que les machines doivent être en mesure de détecter les émotions pour mieux s'adapter aux humains qui les utilisent. Effectivement, il paraît important de les reconnaître pour pouvoir ajuster les réponses d'un système informatique, lorsque l'utilisateur est sujet à des sentiments de frustration. D'autant plus, lorsque la distance entre conseiller et client grandit et davantage encore lorsque l'utilisateur est confronté à un dispositif de réponses automatiques.

Modélisation des émotions

Détecter des émotions est un problème complexe. Il faut d'abord parvenir à définir ce que sont les émotions et comment les modéliser. Hudlicka (Hudlicka, 2011) en propose la définition suivante :

« [...] Nous pouvons considérer les émotions comme des états qui reflètent le jugement évaluatif de l'environnement, de soi-même et des autres agents sociaux, à la lumière des buts et croyances de l'organisme, qui motivent et coordonnent le comportement adaptatif ».

Outre l'approche de la théorie de l'*appraisal* (Scherer, 2005) et l'approche discrète (Ekman, 1992), il existe l'approche dimensionnelle, développée par Russel. Cette approche consiste à

différencier les émotions dans un espace dimensionnel. Ainsi, le circumplex de Russel se compose de deux dimensions principales : la *valence* et l'*arousal* qui se définissent comme étant respectivement, l'aspect positif ou négatif d'une situation et la mobilisation d'énergie (Russell, 1980). On peut trouver des modèles avec des dimensions supplémentaires telles que la dominance notamment dans le modèle PAD (Plaisir, Activation et Dominance) proposé par Mehrabian (Mehrabian, 1996).

Le choix d'un modèle est en grande partie dépendant de l'usage, du niveau de granularité nécessaire et des données disponibles. Ainsi, les termes de « modèles d'émotions » peuvent être considérés comme trop vagues pour donner une idée concrète de ce qu'ils représentent (Hudlicka, 2011).

Dans le cadre de notre étude, nous avons choisi de commencer avec un modèle prenant en compte les dimensions de la *valence* et de l'*arousal* en utilisant comme référence le circumplex de Russel. Le traitement des données, obtenues lors d'une première expérience, en sera ainsi facilité et leur étude nous permettra ensuite de considérer l'ajout de nouvelles dimensions. De plus, les dimensions de la *valence* et de l'*arousal* sont fortement répandues dans la littérature sur des études empiriques (Olsen, 2016), (Coutrix, 2012), (Ghosh, Ganguly, & Mitra, 2017). Ce modèle est donc robuste et correspond au domaine de la relation client.

Les données comportementales des smartphones

Les smartphones sont en mesure de produire des données de plus en plus détaillées sur les utilisateurs et leur comportement. De ce fait, nous avons choisi de nous appuyer sur ce dispositif pour notre recueil de données que nous listons ici :

1) L'accéléromètre et le gyroscope.

Il s'agit de capteurs qui permettent de récupérer, respectivement, les variations de vitesse du téléphone sur les axes x, y et z et la vitesse angulaire de l'appareil.

L'accéléromètre a déjà été exploité pour la détection d'états psychologiques ou émotionnels. Notamment en 2016, Cui, Li, Zhang, & Zhang (Cui, Li, Zhang, & Zhang, 2016) ont utilisé les données de smartphones pour définir un modèle de détection des émotions. Par induction d'émotions, les chercheurs ont enregistré et filtré des données pour réaliser un modèle d'apprentissage automatique des émotions colère et joie, avec une précision de plus de 80%. De la même manière, en 2016, Olsen a effectué une analyse des émotions à partir de l'accéléromètre et du gyroscope, en prenant cette fois comme modèle les dimensions *valence* et *arousal*. (Olsen, 2016).

2) Les coordonnées d'écran et la vitesse de frappe au clavier.

Il est aujourd'hui possible d'enregistrer les coordonnées des points de contact et donc les tracés effectués sur l'écran, ainsi que la dynamique de frappe au clavier. Une étude de Mandra et Coutrix utilise ces coordonnées pour faire le lien entre les gestes et les émotions. (Coutrix, 2012). Les auteurs invitent les participants à réaliser des gestes sur le téléphone et à s'autoévaluer sur les dimensions du modèle PAD, cité précédemment (Mehrabian, 1996). Concernant la saisie au clavier, elle a la particularité d'être sensiblement distincte d'un individu à un autre. En effet, en 2017 une étude portant également sur la détection des émotions utilisait l'interaction des participants sur un smartphone et plus particulièrement leur saisie de texte au clavier tactile (Ghosh, Ganguly, & Mitra, 2017). L'activité est catégorisée par le participant selon une évaluation de son propre état émotionnel (Neutre, Heureux, Triste, Stressé, Relaxé). Par cette méthode les chercheurs ont obtenu une précision de 73% pour prédire l'état émotionnel des participants dans les catégories citées précédemment.

Traitement automatique du langage d'une conversation

Le langage nous paraît une source adaptée à notre étude où des participants vont interagir avec une application de chat. Les sentiments pouvant être recueillis à travers l'expression seront ainsi analysés. La problématique de la mesure des émotions a déjà été étudiée dans un contexte de discussion avec un agent conversationnel par Ma (Ma C. a., 2005). Pour détecter des émotions on peut exploiter : des mots clés considérés comme affectifs en utilisant des bases de données dédiées (Valitutti & Strapparava, 2004), des relations sémantiques, des éléments contextuels, ainsi que des éléments de style ou visuels : ponctuation, émoticon... (Lu, Hsu, Peng, & Chung, 2010).

Utilisation des données comportementales :

Nous considérons que la multimodalité sera un avantage essentiel afin d'obtenir des performances de prédiction satisfaisantes. D'après Olsen (Olsen, 2016) et Coutrix (Coutrix, 2012), certains composants de l'accélération sont corrélés à l'*arousal*, telle que, par exemple, l'accélération moyenne. Nous envisageons donc de trouver des corrélations entre l'évolution de ces données de l'accéléromètre et des émotions induites. De même pour le gyroscope nous formulons l'hypothèse d'une relation entre l'évolution des composants de ce signal et de l'*arousal*. En ce qui concerne la saisie au clavier, nous envisageons un lien significatif entre certaines composantes (vitesse, variance...) et la *valence* et l'*arousal*, tel que présenté dans l'étude (Ghosh, Ganguly, & Mitra, 2017). Nous envisageons une corrélation pour la *valence* et l'*arousal* et les mots utilisés par les participants sur l'application (Ma, Xu, Bai, & Sun, 2012).

IDEE DE PROTOCOLE EXPERIMENTAL

Description générale de l'expérience

On peut distinguer deux types d'études portant sur la mesure d'émotions avec smartphone : les études observationnelles ou les études en laboratoire. La première implique généralement un protocole où les sujets doivent utiliser un smartphone dans leur quotidien pendant une longue durée avec une auto labélisation régulière des données (Ghosh, Ganguly, & Mitra, 2017), (Ma, Xu, Bai, & Sun, 2012). Dans le deuxième type d'étude, en laboratoire, le protocole implique généralement une induction des émotions et une tâche à réaliser sur le smartphone dans un environnement contrôlé (Coutrix, 2012), (Cui, Li, Zhang, & Zhang, 2016).

Pour notre part nous souhaitons mesurer des émotions au cours d'une interaction d'une dizaine de minutes sur une application d'agent conversationnel. Nous envisageons donc un modèle expérimental en laboratoire dans un environnement contrôlé. Comme décrit dans l'état de l'art, les variables que nous souhaitons analyser sont :

- Les mesures de l'accéléromètre et du gyroscope,
- La dynamique de saisie au clavier,
- Les corpus de texte générés par les participants.

Afin de pouvoir labéliser les données recueillies nous allons induire des émotions par des tâches ludiques que les participants devront réaliser en interaction avec un chatbot. Nous souhaitons ainsi générer un engagement cognitif. Nous projetons ensuite de diffuser l'application vers un large public afin de recueillir un volume de données suffisant pour obtenir une plus grande variabilité.

Induire des émotions

Plusieurs auteurs ont montré la possibilité d'induire des états affectifs à partir de matériels tels que des images (Kurdi & Lozano, 2017) des vidéos (Ray, 2007) ou des sons. Ces stimuli ont été testés, avec notamment le modèle d'émotions de

Russel (Russell, 1980), sur un grand nombre de participants. Dans le cadre de notre expérience, utilisant des situations de jeux (énigme, jeux d'attention, test de mémoire), la réussite ou l'échec sont autant de facteurs susceptibles de provoquer une réaction émotionnelle. Notre objectif est d'induire des émotions en parcourant l'espace du modèle de Russel, ce qui peut être simplifié en 5 catégories :

- neutre
- *valence* positive et *arousal* négatif
- *valence* positive et *arousal* positif
- *valence* négative et *arousal* négatif
- *valence* négative et *arousal* positif

Mesures physiologiques et expression faciale

Afin de tester l'efficacité de l'induction, nous allons utiliser des mesures physiologiques objectives. Les technologies développées par Noldus, par exemple, permettent de mesurer les expressions faciales sur les dimensions *valence* et *arousal* (Košir, 2007), (Noldus, 2019). Les mesures de la conductance de la peau, de la fréquence cardiaque ou de la pupillométrie ont été testées sur ce même modèle et nous permettent de vérifier le bon fonctionnement de notre application (Koelstra, et al., 2012).

CONCLUSION ET MISE EN PERSPECTIVE

Bien que les perspectives en termes d'amélioration de services clients soient importantes, notre approche n'a pas encore été validée par des résultats. Il s'agira de prédire avec une précision suffisante, les variations d'émotions par rapport au modèle choisi. En outre, l'engagement des utilisateurs devra être assez fort pour provoquer des émotions intenses et détectables. Fondamentalement, si le protocole permet un contrôle correct des conditions, il limite la généralisation des résultats par rapport au contexte d'un service client où l'interaction peut se passer avec un humain au lieu du chatbot.

Comme évoqué en introduction, nous envisageons plusieurs perspectives à ces travaux dans le domaine de la relation clients. Lorsque des agents conversationnels ont vocation à répondre à des demandes complexes d'utilisateurs, la détection des émotions est nécessaire pour adapter au mieux la communication avec l'agent conversationnel et pour pouvoir déterminer l'échec d'un échange. En détectant de la colère ou de la frustration, on permet une reprise en main par un humain, capable de traiter les demandes les plus complexes et les cas particuliers. Nous envisageons aussi d'autres perspectives pour détecter des situations conflictuelles pour les conseillers eux-même, avec la possibilité de les assister dans la gestion de ces situations éprouvantes.

Au-delà de la performance, facilement mesurable par les indicateurs de production, la mise en place d'outils permettant de mesurer les apports, pour l'humain, de ces nouveaux outils nous semblent une avancée pour la pratique ergonomique. Ainsi, le coût des adaptations sans cesse nécessaires pour les utilisateurs, dans et hors du contexte du travail pourrait être mieux appréhendé.

Ainsi, la question de l'impact général de l'automatisation et de l'intégration globale des technologies dans les métiers du service client mérite d'être posée et les outils de détection des émotions peuvent y apporter un début de réponse.

BIBLIOGRAPHIE

- Brave, S. a. (2003). Emotion in human-computer interaction. *The human-computer interaction handbook: fundamentals, evolving technologies and emerging applications*, 81-96.
- Coutrix, C. &. (2012). Identifying emotions expressed by mobile users through 2D surface and 3D motion gestures. *ACM Conference on Ubiquitous Computing*, 311-320.

- Cui, L., Li, S., Zhang, W., & Zhang, Z. &. (2016). Emotion detection from natural walking. *International Conference on Human Centered Computing*, 23-33.
- Ekman. (1992). An argument for basic emotions. *Cognition & emotion*, 169-200.
- Ghosh, S., Ganguly, N., & Mitra, B. &. (2017). Evaluating effectiveness of smartphone typing as an indicator of user emotion. *Affective Computing and Intelligent Interaction (ACII)*, 146-151.
- Hudlicka, E. (2011). Guidelines for designing computational models of emotions. *International Journal of Synthetic Emotions (IJSE)*, IGI Global,, 26-79.
- Koelstra, S., Muhl, C., Soleymani, M., Lee, J.-S., Yazdani, A., Ebrahimi, T., et al. (2012). Deap: A database for emotion analysis; using physiological signals. *IEEE Transactions on Affective Computing, IEEE*., 18-31.
- Košir, A. &. (2007). Emotion Elicitation in a Socially Intelligent Service: The Typing Tutor. *Computers, Multidisciplinary Digital Publishing Institute*, 14.
- Kurdi, B., & Lozano, S. &. (2017). Introducing the Open Affective Standardized Image Set (OASIS). *Behavior Research Methods*, 457-470.
- Lu, C.-Y., Hsu, W. W., Peng, H.-T., & Chung, J.-M. &.-M. (2010). Emotion sensing for internet chatting: a web mining approach for affective categorization of events. *Computational Science and Engineering (CSE)*, 2010 *IEEE 13th International Conference* , 295-301.
- Ma, C. a. (2005). A chat system based on emotion estimation from text and embodied conversational messengers}. *Active Media Technology*, 546 548.
- Ma, Y., Xu, B., Bai, Y., & Sun, G. &. (2012). Daily mood assessment based on mobile phone sensing. *Wearable and implantable body sensor networks (BSN)*, 2012 *ninth international conference on*., 142-147.
- Maslow, A. H. (1968). *Toward a psychology of being*. Simon and Schuster.
- Mehrabian, A. (1996). Pleasure-arousal-dominance: A general framework for describing and measuring individual differences in temperament. *Current Psychology vol.14*, p 261-292.
- Noldus, I. T. (2019). *Noldus FaceReader*., Retrieved 02 12, 2019, from <http://www.noldus.com/>
- Olsen, A. F. (2016). Smartphone accelerometer data used for detecting human emotions. *Systems and Informatics (ICSAI)*, 2016 *3rd International Conference on*., 410-415.
- Picard, R. W. (1995). *Affective computing*.
- Ray, R. D. (2007). Emotion elicitation using films. *Handbook of emotion elicitation and assessment*.
- Russell, J. A. (1980). A circumplex model of affect. *Journal of personality and social psychology vol 39*, p. 1161.
- Scherer, K. R. (2005). What are emotion? and how can they be measured? *Social science information*, 695-729.
- Valitutti, A., & Strapparava, C. &. (2004). Developing affective lexical resources. *PsychNology Journal*., 61-83.
- Velkovska, J. &. (2014). Parler aux machines, coproduire un service. Intelligence artificielle et travail du client dans les services vocaux automatisés. *E. Kessous et A. Mallard, La fabrique de la vente. Le travail commercial dans les télécommunications*, Presses des Mines, 97-128.